

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-200574

(43)Date of publication of application : 31.07.1998

(51)Int.Cl.

H04L 12/56

(21)Application number : 09-004720

(71)Applicant : DIGITAL VISION LAB:KK

(22)Date of filing : 14.01.1997

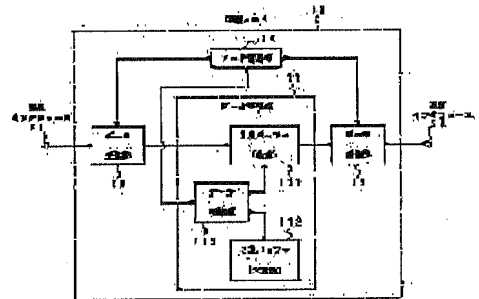
(72)Inventor : TANI HIDEAKI

## (54) RELAY NODE SYSTEM AND RELAY CONTROL METHOD IN THE SYSTEM

## (57)Abstract:

PROBLEM TO BE SOLVED: To temporarily store large capacity burst data, without bringing an overflow which is capable of reducing delay attended with relay of a stream regardless of its low cost.

SOLUTION: A data reception part 12 receives a data stream, transmitted from a transmission node and data in the stream are written in a primary buffer 111. A data transmission part 13 reads data stored in the primary buffer 111 and sends the data to a reception node. A data moving part 113 saves the succeeding data from the primary buffer 111 to a secondary buffer 112 in units of blocks, when the possibility of overflow occurs at an input side of the primary buffer 111. The data in the secondary buffer 112 are moved to the primary buffer 111 from the secondary buffer 112 by the data moving part 113, in accordance with timing transmitted continuously by the data transmission part 13 following the data stored in the primary buffer 111 which is not an object of save processing.



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-200574

(43) 公開日 平成10年(1998) 7月31日

(51) Int.Cl.<sup>4</sup>

H 0 4 L 12/56

識別記号

F I

H 0 4 L 11/20

1 0 2 B

審査請求 有 請求項の数18 O L (全 18 頁)

(21) 出願番号 特願平9-4720  
 (22) 出願日 平成9年(1997) 1月14日

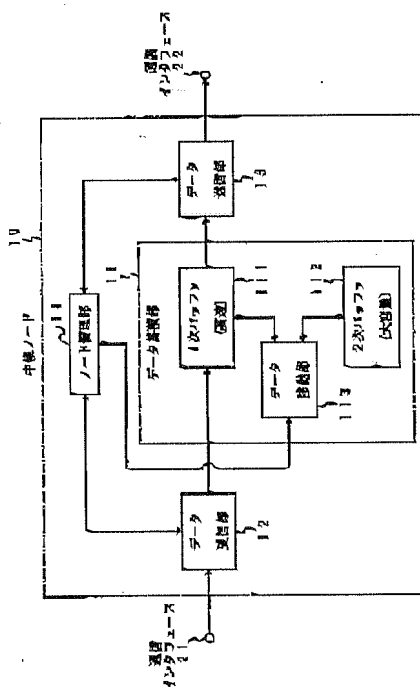
(71) 出願人 396001360  
 株式会社デジタル・ビジョン・ラボラトリーズ  
 東京都港区赤坂七丁目3番37号  
 (72) 発明者 谷 英明  
 東京都港区赤坂七丁目3番37号 株式会社  
 デジタル・ビジョン・ラボラトリーズ内  
 (74) 代理人 弁理士 鈴江 武彦 (外5名)

(54) 【発明の名称】 中継ノードシステム及び同システムにおける中継制御方法

## (57) 【要約】

【課題】 低価格でありながら、ストリーム中継に伴う遅延を小さくでき、大容量のバースト的なデータをオーバーフローを招かずに一時蓄積できるようにする。

【解決手段】 送信ノードから送信されたデータストリームをデータ受信部12にて受信して、そのストリーム中のデータを1次バッファ111に書き込む。データ送信部13は1次バッファ111内の蓄積データを読み出して受信ノードへ送出する。データ移動部113は、1次バッファ111の入力側のオーバーフローの危険性が生じた場合に、それ以降のデータについて1次バッファ111から2次バッファ112へブロック単位で退避する。2次バッファ112内のデータは、退避処理の対象とならなかった1次バッファ111内の蓄積データに続いてデータ送信部113により連続的に送出可能なタイミングに合わせて、データ移動部113により2次バッファ112から1次バッファ111に移動される。



(2)

特開平10-200574

1

2

## 【特許請求の範囲】

【請求項1】 送信ノードから送信された一連のデータ要素からなるデータストリームを受信するデータ受信手段と、

前記データ受信手段により受信されたデータストリーム中のデータを一時蓄積するための高速メモリから構成される1次バッファ手段と、

前記1次バッファ手段に蓄積されたデータを取り出して受信ノードに送出するデータ送信手段と、

前記1次バッファ手段の記憶領域を補充するための当該1次バッファ手段より低速大容量の2次バッファ手段と、

前記1次バッファ手段の入力側のオーバフローの危険性を検出して、前記1次バッファ手段から前記2次バッファ手段へブロック単位でデータを移動する退避処理を行うと共に、前記2次バッファ手段に退避されたデータを前記データ送信手段によるデータ送出のために前記1次バッファ手段へブロック単位で移動する復帰処理を行うデータ移動手段とを具備することを特徴とする中継ノードシステム。

【請求項2】 前記データ移動手段は、中継の対象となる前記各データストリーム毎に、前記退避処理の単位となるブロック長 $b_j$ と、そのストリームのデータを退避なしに前記1次バッファ手段に一時蓄積するのに必要な領域のサイズである1次バッファ割り当てサイズ $B_j$ とを含むパラメータを設定するパラメータ設定手段を有しており、

前記パラメータ設定手段は、対応する前記データストリームの入力レート $R_{in,j}$ が出力レート $R_{out,j}$ より大きい場合には、前記ブロック長 $b_j$ だけのデータの先頭部が前記1次バッファ手段に到来してから、当該データが前記1次バッファ手段及び前記2次バッファ手段間を往復し、前記1次バッファ手段から全て送出されるまでに要すると推定される時間 $T_j$ の間に、該当する前記データストリームについて前記データ送信手段により前記1次バッファ手段から送出されと推定されるデータ量以上の値に前記1次バッファ割り当てサイズ $B_j$ を設定することを特徴とする請求項1記載の中継ノードシステム。

【請求項3】 前記データ移動手段は、前記退避処理が行われていない前記各データストリーム毎に、そのストリームについての前記1次バッファ手段内のデータ蓄積量を監視し、前記データ蓄積量が対応する前記データストリームに関して前記時間 $T_j$ の間に送出されると推定されるデータ量を超えた場合に前記入力側オーバフローの危険性を検出し、当該データストリームについて前記退避処理を伴わない退避なしモードから前記退避処理を伴う退避ありモードに切り替え設定するモード設定手段を有している請求項2記載の中継ノードシステム。

【請求項4】 前記データ移動手段は、前記退避ありモードが設定されている前記データストリーム中のデータ

が対応する前記ブロック長 $b_j$ だけ前記1次バッファ手段に書き込まれる毎に、そのブロック長 $b_j$ のデータを前記2次バッファ手段に移動する前記退避処理を行う退避手段と、前記退避手段により前記ブロック長 $b_j$ のデータが前記2次バッファ手段に移動される毎に、当該データの先頭部が前記データ送信手段により送出される時刻を推定して目標復帰完了時刻として設定し、当該データを前記2次バッファ手段から前記1次バッファ手段に移動するのに必要な時間だけ前記設定した目標復帰完了時刻より遡った時刻より当該データを前記2次バッファ手段から前記1次バッファ手段に移動する前記復帰処理を開始する復帰手段とを有していることを特徴とする請求項3記載の中継ノードシステム。

【請求項5】 前記モード設定手段は、前記復帰手段による前記復帰処理の完了時において対応するデータストリームについての後続の復帰対象データが所定数以内である状態が、所定回数連続した場合には、当該データストリームについて前記退避ありモードから前記退避なしモードに切り替え設定することを特徴とする請求項4記載の中継ノードシステム。

【請求項6】 前記パラメータ設定手段は、前記退避なしモードが設定されている前記各データストリームの各々に対して割り当てられている前記1次バッファ割り当てサイズ $B_j$ 、前記退避ありモードが設定されている前記各データストリームにより前記1次バッファ手段上で占められるサイズ、及び前記1次バッファ手段のサイズをもとに、前記1次バッファ手段の利用率 $\rho_1$ を求めると共に、前記退避ありモードが設定されている前記各データストリームに関する前記ブロック長 $b_j$ のデータの退避処理でのトラフィック及び復帰処理でのトラフィックをもとに、前記2次バッファ手段に対するアクセスチャネル利用率 $\rho_2$ を求め、前記利用率 $\rho_1$ 及び $\rho_2$ のいずれか一方が1を超えている場合には、少なくとも1つの前記データストリームについて前記ブロック長 $b_j$ を変更して前記利用率 $\rho_1$ 及び $\rho_2$ を再計算することを特徴とする請求項2記載の中継ノードシステム。

【請求項7】 前記パラメータ設定手段は、前記利用率 $\rho_1$ が1を超えている場合には、少なくとも1つの前記データストリームについて前記ブロック長 $b_j$ を減少する方向に変更し、前記利用率 $\rho_2$ が1を超えている場合には、少なくとも1つの前記データストリームについて前記ブロック長 $b_j$ を増加する方向に変更することを特徴とする請求項6記載の中継ノードシステム。

【請求項8】 前記パラメータ設定手段は、前記退避なしモードが設定されている前記各データストリームの各々に対して割り当てられている前記1次バッファ割り当てサイズ $B_j$ の対応データストリームの入力変動を吸収するための第1のマージン値 $\alpha_{1,j}$  ( $0 \leq \alpha_{1,j} < 1$ )を考慮した総和 $\Sigma 1$ と、前記退避ありモードが設定されている前記各データストリームにより前記1次バッファ手段上

(3)

特開平10-200574

3

4

で占められるサイズの対応データストリームの退避並びに復帰処理時間の変動を吸収するための第2のマージン値 $\alpha 2j$  ( $0 \leq \alpha 2j < 1$ ) を考慮した総和 $\Sigma 2$ と、前記1次バッファ手段のサイズ $L1$ とから、前記利用率 $\rho 1$ を $\rho 1 = (\Sigma 1 + \Sigma 2) / L1$

の演算により求めると共に、前記退避ありモードが設定されている前記各データストリームに関する前記ブロック長 $b_j$ のデータの退避処理でのトラフィック及び復帰処理でのトラフィックをもとに、前記2次バッファ手段に対するアクセスチャネル利用率 $\rho 2$ を求め、前記利用率 $\rho 1$ 及び $\rho 2$ のいずれか一方が1を超えている場合には、少なくとも1つの前記データストリームについて前記ブロック長 $b_j$ を変更して前記利用率 $\rho 1$ 及び $\rho 2$ を再計算することを特徴とする請求項2記載の中継ノードシステム。

【請求項9】 前記パラメータ設定手段は、前記利用率 $\rho 1$ が1を超えている場合には、少なくとも1つの前記データストリームについて前記ブロック長 $b_j$ を減少する方向に変更し、前記利用率 $\rho 2$ が1を超えている場合には、少なくとも1つの前記データストリームについて前記ブロック長 $b_j$ を増加する方向に変更し、前記利用率 $\rho 1$ 及び $\rho 2$ が共に1を超えている場合には、前記第1のマージン値 $\alpha 1j$ または第2のマージン値 $\alpha 2j$ の少なくとも一方を減少する方向に変更して前記利用率 $\rho 1$ 及び $\rho 2$ を再計算することを特徴とする請求項8記載の中継ノードシステム。

【請求項10】 送信ノードから送信された一連のデータ要素からなるデータストリームを受信して、そのデータを受信ノードへ再送出する中継ノードシステムにおける中継制御方法であって、前記送信ノードから送信された前記データストリームを受信して、そのストリーム中のデータを高速メモリから構成される1次バッファ手段に書き込み、通常はその1次バッファ手段に書き込んだデータを該当するデータストリームに要求される出力レートに応じて読み出して受信ノードへ送出し、前記1次バッファ手段の入力側のオーバーフローの危険性が生じた場合には、それ以降のデータについて前記1次バッファ手段から当該1次バッファ手段の記憶領域を補完するための当該1次バッファ手段より低速大容量の2次バッファ手段へブロック単位で移動する退避処理を行い、前記退避処理によって前記2次バッファ手段に退避されたデータを、前記退避処理の対象とならなかった前記1次バッファ手段内の蓄積データに続いて連続的に前記受信ノードへ送出可能なように、前記2次バッファ手段から前記1次バッファ手段へブロック単位で移動する復帰処理を行うことを特徴とする中継ノードシステムにおける中継制御方法。

【請求項11】 中継の対象となる前記各データストリ

10

20

30

40

50

ーム毎に、前記退避処理の単位となるブロック長 $b_j$ と、そのストリームのデータを退避なしに前記1次バッファ手段に一時蓄積するのに必要な領域のサイズである1次バッファ割り当てサイズ $B_j$ とを含むパラメータを設定し、対応する前記データストリームの入力レート $R_{in j}$ が出力レート $R_{out j}$ より大きい場合には、前記ブロック長 $b_j$ だけのデータの先頭部が前記1次バッファ手段に到来してから、当該データが前記1次バッファ手段及び前記2次バッファ手段間を往復し、前記1次バッファ手段から全て送出されるまでに要すると推定される時間 $T_j$ の間に、該当する前記データストリームについて前記1次バッファ手段から送出されと推定されるデータ量以上の値に前記1次バッファ割り当てサイズ $B_j$ を設定することを特徴とする請求項10記載の中継ノードシステムにおける制御方法。

【請求項12】 前記退避処理が行われていない前記各データストリーム毎に、そのストリームについての前記1次バッファ手段内のデータ蓄積量を監視し、前記データ蓄積量が対応する前記データストリームに関して前記時間 $T_j$ の間に送出されると推定されるデータ量を超えた場合に前記入力側オーバーフローの危険性を検出し、当該データストリームについて前記退避処理を伴わない退避なしモードから前記退避処理を伴う退避ありモードに切り替え設定する請求項11記載の中継ノードシステムにおける中継制御方法。

【請求項13】 前記退避ありモードが設定されている前記データストリーム中のデータが対応する前記ブロック長 $b_j$ だけ前記1次バッファ手段に書き込まれる毎に、そのブロック長 $b_j$ のデータを前記2次バッファ手段に移動する前記退避処理を行うと共に、当該データの先頭部が前記受信ノードへ送出される時刻を推定して目標復帰完了時刻として設定し、当該データを前記2次バッファ手段から前記1次バッファ手段に移動するのに必要な時間だけ前記設定した目標復帰完了時刻より溯った時刻より当該データを前記2次バッファ手段から前記1次バッファ手段に移動する前記復帰処理を開始することを特徴とする請求項12記載の中継ノードシステムにおける中継制御方法。

【請求項14】 前記復帰処理の完了時において対応するデータストリームについての後続の復帰対象データが所定数以内である状態が、所定回数連続した場合には、当該データストリームについて前記退避ありモードから前記退避なしモードに切り替え設定することを特徴とする請求項13記載の中継ノードシステムにおける中継制御方法。

【請求項15】 前記退避なしモードが設定されている前記各データストリームの各々に対して割り当てられている前記1次バッファ割り当てサイズ $B_j$ 、前記退避ありモードが設定されている前記各データストリームにより前記1次バッファ手段上で占められるサイズ、及び前

(4)

特開平10-200574

5

6

記1次バッファ手段のサイズをもとに、前記1次バッファ手段の利用効率 $\rho_1$ を求めると共に、前記回避ありモードが設定されている前記各データストリームに関する前記ブロック長 $b_j$ のデータの回避処理でのトラフィック及び復帰処理でのトラフィックをもとに、前記2次バッファ手段に対するアクセスチャネル利用率 $\rho_2$ を求め、前記利用率 $\rho_1$ 及び $\rho_2$ のいずれか一方が1を超えている場合には、少なくとも1つの前記データストリームについて前記ブロック長 $b_j$ を変更して前記利用率 $\rho_1$ 及び $\rho_2$ を再計算することを特徴とする請求項14記載の中継ノードシステムにおける中継制御方法。

【請求項16】 前記回避なしモードが設定されている前記各データストリームの各々に対して割り当てられている前記1次バッファ割り当てサイズ $B_j$ の対応データストリームの入力変動を吸収するための第1のマージン値 $\alpha_1j$  ( $0 \leq \alpha_1j < 1$ )を考慮した総和 $\Sigma 1$ と、前記回避ありモードが設定されている前記各データストリームにより前記1次バッファ手段上で占められるサイズの対応データストリームの回避並びに復帰処理時間の変動を吸収するための第2のマージン値 $\alpha_2j$  ( $0 \leq \alpha_2j < 1$ )を考慮した総和 $\Sigma 2$ と、前記1次バッファ手段のサイズ $L_1$ とから、前記利用率 $\rho_1$ を

$$\rho_1 = (\Sigma 1 + \Sigma 2) / L_1$$

の演算により求めると共に、前記回避ありモードが設定されている前記各データストリームに関する前記ブロック長 $b_j$ のデータの回避処理でのトラフィック及び復帰処理でのトラフィックをもとに、前記2次バッファ手段に対するアクセスチャネル利用率 $\rho_2$ を求め、前記利用率 $\rho_1$ 及び $\rho_2$ のいずれか一方が1を超えている場合には、少なくとも1つの前記データストリームについて前記ブロック長 $b_j$ を変更して前記利用率 $\rho_1$ 及び $\rho_2$ を再計算することを特徴とする請求項14記載の中継ノードシステムにおける中継制御方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、一連のデータ要素からなるビデオデータ等のデータストリームを送信ノードから受信ノードに伝送する通信システムに係り、特にネットワーク上において、送信ノードからデータストリームを受信し、そのデータを受信ノードへ再送出する中継ノードシステム及び同システムにおける中継制御方法に関する。

【0002】

【従来の技術】

(a1) 中継ノードの基本構成

送信ノードから一連のデータ要素からなるデータストリームを受信し、そのデータを受信ノードへ再送出する中継ノードは、図11のように構成されるのが一般的である。

【0003】 図11において、中継ノード100は、デ

ータ蓄積部101と、通信インタフェース110を介してネットワークからデータストリームを構成する連続する一連のデータ要素であるデータパケット（或いはデータフレーム）を受信しこのパケットに含まれるデータをデータ蓄積部101に書き込む機能を持つデータ受信部102と、データ蓄積部101からデータを読み出し対応するデータストリームのデータパケットに組み立てて通信インタフェース120を介してネットワークへ送出する機能を持つデータ送信部103と、データ蓄積部101、データ受信部102及びデータ送信部103のそれぞれに対して状態の獲得や動作の制御を司るノード管理部104とから構成されるのが一般的である。

【0004】 図11の構成の中継ノード100を備えた通信システムでは、送信ノードからデータストリーム、即ち、連続する一連のデータパケットが当該中継ノード100宛に送出されると、これらのパケットは当該中継ノード100内のデータ受信部102によってデータ蓄積部101に蓄えられる一方、データ送信部103によって再び読み出され、後段の中継ノードまたは受信端末へと再送出される。

(a2) 中継ノード挿入の目的

図11に示したような構成の中継ノード100を送受信ノード間に挿入する目的には、一般的に網間接続、アドレス変換、データフォーマット変換、パケット転送間隔の再構成、指定されたルールに基づく選択的パケット廃棄などが含まれる。特に、ビデオデータなどのリアルタイム情報を含むデータストリームの中継においては、安定したストリームレート制御能力の不足する送信ノードからバースト的にデータを受信し、このデータを高精度なレート制御に基づき再送出する中継処理が有効であり、大規模なジッタを吸収できることが要求される。

(a3) 中継ノードでのデータ蓄積方式

図11に示した中継ノード100内のデータ蓄積部102には、従来は、次の2つのデータ蓄積方式のうちのいずれかの形態が適用されるのが一般的である。

【0005】 第1は、到着データを半導体バッファメモリに一時蓄積するデータ蓄積方式であり、第2は、到着データを磁気ディスクなどの2次大容量記憶媒体に一時蓄積するデータ蓄積方式である。

【0006】 ところが、上記第1のデータ蓄積方式は、バッファ管理が高速に行え、入出力間の遅延が小さいという利点を持つ反面、より大きなジッタ吸収のために蓄積容量を拡大するには高価格となるという欠点を持つ。一方、上記第2のデータ蓄積方式は、大容量蓄積には適しているが、記憶媒体書き込み/読み出しに伴う遅延時間及びその変動が大きいという欠点を持つ。

【0007】

【発明が解決しようとする課題】 データストリームの中継する中継ノードには、大規模なジッタを吸収することが要求され、そのためには、中継データの蓄積容量の大

(5)

特開平10-200574

7

容量化と低遅延の中継動作の両方が必要となる。

【0008】しかしながら従来技術では、上記した2つのデータ蓄積方式のうちの第1のデータ蓄積方式では、蓄積容量の大規模化がコスト的に難しく、第2のデータ蓄積方式では、低遅延中継処理の実現が難しいという問題があった。また、中継ノードは一般に同時に複数のデータストリームの中継を取り扱う必要があることから、メモリなどのノード内共有資源を効率的に利用することも要求されるが、従来技術では、このような要求に応えられなかった。

【0009】本発明は上記事情を考慮してなされたものでその目的は、比較的低価格でありながら、ストリーム中継に伴う遅延を小さくできると共に、大容量のバースト的なデータをオーバフローの発生を招くことなく一時蓄積することができる中継ノードシステム及び同システムにおける制御方法を提供することにある。

【0010】

【課題を解決するための手段】本発明の中継ノードシステムは、送信ノードから送信された一連のデータ要素からなるデータストリームを受信するデータ受信手段と、このデータ受信手段により受信されたデータストリーム中のデータを一時蓄積するための高速メモリから構成される1次バッファ手段と、この1次バッファ手段に蓄積されたデータを取り出して受信ノードに送出するデータ送信手段と、上記1次バッファ手段の記憶領域を補充するための当該1次バッファ手段より低速大容量の2次バッファ手段と、上記1次バッファ手段の入力側のオーバフローの危険性を検出して、当該1次バッファ手段から上記2次バッファ手段へブロック単位でデータを移動する退避処理を行うと共に、上記2次バッファ手段に退避されたデータを上記データ送信手段によるデータ送出のために上記1次バッファ手段へブロック単位で移動する復帰処理を行うデータ移動手段とを備えたことを特徴とする。

【0011】このような構成の中継システムにおいては、中継するデータストリームを一時蓄積する1次バッファ手段を高速メモリで構成していることから、低遅延中継処理が実現できるだけでなく、当該1次バッファ手段の容量を低価格化のために小容量としても、1次バッファ手段の入力側オーバフローの危険性がある場合には、1次バッファ手段に入力されたデータが低速ではあるものの大容量の2次バッファ手段にブロック単位で移動・退避されることから、中継データの蓄積容量の大容量化も実現できる。しかも、2次バッファ手段に退避されたデータが退避処理の対象とならなかった1次バッファ手段内の蓄積データに続いて連続的に受信ノードに送出可能なように、2次バッファ手段から1次バッファ手段へブロック単位でデータを移動する復帰処理が行われることから、出力側のアンダーフローを極力防止できる。

8

【0012】特に、上記データ手段に、中継の対象となる各データストリーム毎に、退避処理の単位となるブロック長 $b_j$ と、そのストリームのデータを退避なしに1次バッファ手段に一時蓄積するのに必要な領域のサイズである1次バッファ割り当てサイズ $B_j$ とを含むパラメータを設定するパラメータ設定手段を持たせ、対応するデータストリームの入力レート $R_{in,j}$ が出力レート $R_{out,j}$ より大きい場合には、上記ブロック長 $b_j$ だけのデータの先頭部が1次バッファ手段に到来してから、当該データが1次バッファ手段及び2次バッファ手段間を往復し、1次バッファ手段から全て送出されるまでに要すると推定される時間 $T_j$ の間に、該当するデータストリームについてデータ送信手段により1次バッファ手段から受信ノードに送出されと推定されるデータ量以上の値に1次バッファ割り当てサイズ $B_j$ を設定するならば、2次バッファ手段に退避したデータを1次バッファ手段に復帰させて1次バッファ手段上の $B_j$ の領域内の蓄積データに続いて連続的に送出することが可能となる。

【0013】また、上記のパラメータ設定により、退避処理が行われていない各データストリーム毎に、そのストリームについての1次バッファ手段内のデータ蓄積量と、対応するデータストリームに関して上記時間 $T_j$ の間に送出されると推定されるデータ量とを比較することで、入力側オーバフローの危険性があるストリームを簡単に検出でき、その際に当該ストリームについて退避処理を伴わない退避なしモードから退避処理を伴う退避ありモードに切り替え設定することにより、入力側オーバフローを一層効果的に抑止できる。

【0014】また、退避ありモードが設定されているデータストリーム中のデータが対応するブロック長 $b_j$ だけ1次バッファ手段に書き込まれる毎に、そのブロック長 $b_j$ のデータを2次バッファ手段に移動する退避処理を行い、ブロック長 $b_j$ のデータが2次バッファ手段に退避される毎に、当該データの先頭部がデータ送信手段により送出される時刻を推定して目標復帰完了時刻として設定し、当該データを2次バッファ手段から1次バッファ手段に移動するのに必要な時間だけ設定した目標復帰完了時刻より溯った時刻より当該データを2次バッファ手段から1次バッファ手段に移動する復帰処理を開始することにより、低速の2次バッファ手段を用いながら、受信ノードへのデータ送出の連続性を一層確保できる。

【0015】また、上記復帰処理の完了時において対応するデータストリームについての後続の復帰対象データが所定数以内である状態が、所定回数連続した場合に、当該データストリームについて退避ありモードから退避なしモードに切り替え設定するとよい。

【0016】また本発明は、退避なしモードが設定されている各データストリームの各々に対して割り当てられている1次バッファ割り当てサイズ $B_j$ 、退避ありモー

50

(5)

特開平10-200574

9

10

ドが設定されている各データストリームにより1次バッファ手段上で占められるサイズ、及び1次バッファ手段のサイズをもとに、1次バッファ手段の利用率 $\rho_1$ を求めると共に、退避ありモードが設定されている各データストリームに関するブロック長 $h_j$ のデータの退避処理でのトラフィック及び復帰処理でのトラフィックをもとに、2次バッファ手段に対するアクセスチャネル利用率 $\rho_2$ を求め、利用率 $\rho_1$ 及び $\rho_2$ のいずれか一方が1を超えている場合には、少なくとも1つのデータストリームについてブロック長 $h_j$ を変更して利用率 $\rho_1$ 及び $\rho_2$ を再計算することをも特徴とする。

【0017】このような構成においては、ブロック長の減少或いは増加操作により、対応するデータストリームによる1次バッファ手段の占有時間を短くしたり、退避及び復帰処理の発生回数を少なくしたりすることができる。

【0018】また、中継対象となる全てのデータストリームの $B_j$ について対応ストリームの入力変動を吸収するための第1のマージン値 $\alpha_1j$  ( $0 \leq \alpha_1j < 1$ )を考慮した総和 $\Sigma 1$ と、退避ありモードが設定されている各データストリームにより1次バッファ手段上で占められる1次バッファ割り当てサイズ $B_j$ 以外のサイズの対応データストリームの退避並びに復帰処理時間の変動を吸収するための第2のマージン値 $\alpha_2j$  ( $0 \leq \alpha_2j < 1$ )を考慮した総和 $\Sigma 2$ との和を1次バッファ手段のサイズ $L_1$ で割った値を $\rho_1$ とすることで、 $\alpha_1j$ 、 $\alpha_2j$ の設定値に応じた対応する変動吸収が可能となる。

【0019】

【発明の実施の形態】以下、本発明の実施の形態につき図面を参照して説明する。

【中継ノードの構成】図1は本発明の一実施形態に係る中継ノード（ストリーム中継装置）10の構成を示すブロック図である。

【0020】図1の中継ノード10は、データ蓄積部11と、通信インタフェース21を介してネットワークからデータストリームを構成する連続する一連のデータパケット（データフレーム）を受信しこのパケットに含まれるデータをデータ蓄積部11に書き込む機能を持つデータ受信部12と、データ蓄積部12からデータを読み出し対応するデータストリームのデータパケットに組み立てて通信インタフェース22を介してネットワークへ送出する機能を持つデータ送信部13と、データ蓄積部11、データ受信部12及びデータ送信部13のそれぞれに対して状態の獲得や動作の制御を司るノード管理部14とから構成される。

【0021】データ蓄積部11は、データ受信部12により受信されたパケット中のデータを一時蓄積するための1次バッファ111と、この1次バッファ111のオーバーフロー発生を回避するように、当該1次バッファ111に不足する記憶領域を補完するための2次バッファ

112と、ノード管理部14の制御のもとでバッファ111、112間のデータ移動を司るデータ移動部113とから構成される。1次バッファ111は、RAM等の半導体メモリを用いて構成される高速小容量バッファであり、2次バッファ112は、磁気ディスク等の大容量記憶媒体を用いて構成される低速大容量バッファである。

【0022】このように本実施形態におけるデータ蓄積部11のデータ一時蓄積機能（を持つ一時蓄積バッファ）は、半導体メモリによる1次バッファ111と大容量記憶媒体による2次バッファ112の組み合わせによって実現される。

【0023】データ蓄積部11内のデータ移動部113は、ノード管理部14の制御のもとで上記両バッファ111、112間のデータ移動に関するスケジューリングを行うようになっている。またデータ移動部113は、ノード管理部14と連携して、バッファ111、112内の蓄積データ量、入力データストリームの予測値、データ到着イベント、データ再送出イベントに基づくアルゴリズムで最適なデータ移動及び再送出を制御するようになっている。[バッファ111、112間のデータ移動制御]ここで、データ蓄積部11におけるバッファ111、112間のデータ移動制御の要点について説明する。

【0024】まず、1次バッファ111と2次バッファ112との間でデータ移動を行う目的は、到着間隔やブロック長に確率的変動のある1次バッファ111へのデータ入力に対し、当該1次バッファ111のオーバーフロー発生を回避することと、後段（の中継ノードまたは受信端末）へのデータ再送出のために1次バッファ111から安定したデータ出力を行えるようにすることにある。なお以下では、1次バッファ111から2次バッファ112へのデータ移動を「退避」、その逆の方向へのデータ移動、即ち2次バッファ112から1次バッファ111へのデータ移動を「復帰」と呼ぶことにする。

【0025】本実施形態におけるデータ蓄積部11の構成では、データの退避処理及び復帰処理において、その処理自身に有限の処理時間（固定的な処理時間）を費やす。また、2次バッファ112を1つまたは少数の大容量記憶媒体により構成するものとする、上記退避処理と復帰処理の間でアクセスの競合が発生する。また複数のストリームが同一の2次バッファ112を利用する場合には、ストリーム間での競合も生じる。そこで、競合の結果生じる待ち合わせ時間に起因する1次バッファ111入力側のオーバーフローや出力側のアンダーフローを極力防止するように、データ移動部113において最適なスケジューリングを行う必要がある。

【0026】上記のスケジューリングの効率は、各ストリームに対するブロック長の設定値によって制御することができる。即ち、退避及び復帰の際の処理単位となる

(7)

特開平10-200574

11

ブロック長（退避ブロック長）を $b$ （単位は、例えばbyte）とすると、直接的には $b$ を小さくするならば、1次バッファ111の占有時間が短くなり、1次バッファ111のオーバーフローの発生を抑制するのに寄与する。しかしその反面、退避及び復帰の回数が多くなることから、退避及び復帰に伴う前述の処理時間固定部分が全体として増加する。その結果、2次バッファ112へのアクセスが相対的に増加し、待ち合わせ遅延によるバッファオーバーフローの発生を増加させる一因となる。このため、上記のスケジューリングにおいては、この点も考慮する必要がある。

【データ移動制御の手順の概略】次に、データ蓄積部11内でのバッファ111、112間のデータ移動制御の手順の概略につき、1つのストリームのデータの退避及び復帰処理を例に図2のタイミングチャートを参照して説明する。

【0027】データ蓄積部11ではまず、1次バッファ111内に入力データを受信するためのサイズ $L_{in}$ （単位は、例えばbyte）の入力バッファ領域を確保して、データ受信部12によるネットワーク上の送信ノードからのデータパケット受信に伴うデータの入力を待つ。ここで $L_{in}$ （入力バッファ領域長）は、1次バッファ111に対するアクセス単位の整数倍で、且つ退避ブロック長 $b$ 以上の値に設定される。

【0028】データストリームの開始時におけるデータ入力の待機時間 $w_{in}$ は一般に予測困難であるが、データストリーム転送途中においては1つ前のデータブロックの最後のタイミングを参考にしてスケジューリングを行うことにより、この待機時間 $w_{in}$ を最小にすることができる。

【0029】次に、1次バッファ111内の入力バッファ領域へのデータ入力が開始される。この入力バッファ領域にブロック長 $b$ だけのデータが入力されるのに要する時間（データ入力時間） $t_{in}$ （の平均値）は、実際の転送データパケット長などに依存するが、付加的な入力バッファを置かない場合、データストリームの平均的な入力レートを $R_{in}$ （単位は通常はbit/secであるが、ここではブロック長 $b$ の単位byteに合わせてbyte/sec）とすると、次式で与えられる。

【0030】 $t_{in}=b/R_{in}$ 次に、ブロック長 $b$ だけのデータ入力が終了した段階で、データ移動部113は、このデータをそのまま1次バッファ111に残すか、或いは2次バッファ112へ退避（移動）するかを判断する。ここでの判断の基準については後述する。また、サイズ $L_{in}$ の入力バッファ領域に $L_{in}$ 分のデータが入力されると、1次バッファ111内にサイズ $L_{in}$ の次の入力バッファ領域が確保される。

【0031】データ移動部113により1次バッファ111に残す判断が与えられたストリームに属するデータは、そのまま1次バッファ111上に蓄積され（蓄積時

12

間を $d$ とする）、データ送信部13による再送出（のための読み出し）を待つ。

【0032】これに対し、あるストリームに対してデータ移動部113が2次バッファ112への退避を判断した場合には、当該データ移動部113はその判断を変更するまでの間、このストリームに属するデータをブロック長 $b$ を単位に全て2次バッファ112へ退避する。この場合、1次バッファ111から2次バッファ112へのデータブロック移動の処理が発生する。2次バッファ112へのデータブロック移動は、データ蓄積部11において同時に扱われるストリームの全てについて行われるため、競合がある場合には待ち合わせ時間（退避待ち時間） $w_1$ が発生する。したがって退避に要する時間（退避時間）は、1次バッファ111から2次バッファ112へ実際にデータブロックを転送して書き込む退避処理に要する時間（退避処理時間）を $t_1$ とすると、 $w_1+t_1$ となる。

【0033】データ移動部113は、2次バッファ112に退避したデータブロックを、そのブロックの再送出のタイミングに間に合うように1次バッファ111に復帰する復帰処理を行わなければならない。そのため、1次バッファ111上にサイズ $L_{out}$ （単位は、例えばbyte）の出力バッファ領域を確保する。ここで $L_{out}$ （出力バッファ領域長）は、1次バッファ111に対するアクセス単位の整数倍で、且つ退避ブロック長 $b$ より大きい値に設定される。

【0034】さて、復帰処理においても、上記退避処理と同様に、待ち合わせ時間（復帰待ち時間） $w_2$ が発生する。したがって復帰に要する時間（復帰時間）は、2次バッファ112から実際にデータブロックを読み出して1次バッファ111へ転送する復帰処理に要する時間（復帰処理時間）を $t_2$ とすると、 $w_2+t_2$ となる。

【0035】上記データブロックの退避処理時間 $t_1$ 及び復帰処理時間 $t_2$ は、2次バッファ112を構成する蓄積デバイスの特性に依存する。例えば磁気ディスクや光ディスクなど回転体の記録媒体を利用するデバイスでは、データブロックの退避処理には、データ書き込みの際にヘッドを移動するシーク時間並びに回転待ち時間（ブロック長に依存せず）と、実際にデータを転送・書き込みする時間（ブロック長に依存）がかかる。またデータブロックの復帰処理の際も同様である。これを一般化すると、退避・復帰処理時間は、ブロック長 $b$ に依存しない部分と、ブロック長 $b$ に依存する部分との和になる。

【0036】最も単純な退避処理時間 $t_1$ 及び復帰処理時間 $t_2$ の推定値の近似式は、退避処理時のブロック長 $b$ に依存しない時間（退避処理時間固定部分）を $f_1$ 、2次バッファ112への書き込み速度を $r_1$ 、復帰処理時のブロック長 $b$ に依存しない時間（復帰処理時間固定部分）を $f_2$ 、2次バッファ112からの読み出し速度



(8)

特開平10-200574

13

を $r2$ とすると、以下のようになる。

【0037】 $t1 = f1 + b / r1$

$t2 = f2 + b / r2$

上記 $t1$ 、 $t2$ は推定値であることから、以下の説明では、それぞれ推定退避処理時間、推定復帰処理時間と呼ぶことにする。

【0038】さて、2次バッファ112から1次バッファ111に復帰されたデータブロックは再送待を待ち、待ち合わせ時間（送出待ち合わせ時間） $wout$ の後、データ送信部13により後段の受信ノード（具体的には中継ノードまたは受信端末）に再送される。この再送に要する時間（データ送出時間） $tout$ は、データストリームの平均的な出力レートを $Rout$ （単位は通常は $bit/sec$ であるが、ここではブロック長 $b$ の単位 $byte$ に合わせて $byte/sec$ ）とすると、次式で与えられる。

【0039】 $tout = b / Rout$

固定ビットレートのビデオストリームを中継する場合などにおいては、上記復帰データの再送のタイミングをノード管理部14にて予め推定できる。したがって、推定した再送のタイミングに合わせてノード管理部14がデータ移動部113による復帰処理を起動させることで、上記送出待ち合わせ時間 $wout$ を最小化することができる。

【0040】本実施形態によるデータ一時蓄積メカニズムの動作効率は、1次バッファ111の容量及び2次バッファ112へのアクセススループットによって制限されるが、これについては後述する。

【中継制御】次に、図1の構成の中継ノード10における上述のデータ移動制御を含む中継制御について説明する。

【0041】まず、（ビデオサーバ、中継ノードなどの）前段の送信ノードから図1の中継ノード10に対してデータストリーム $j$ の転送を要求する際には、データ入力レート $Rinj$ が通知（申告）される。ノード管理部14は、送信ノードから通知されたデータ入力レート $Rinj$ を内部保持した後、データ移動部113と共同して、データストリーム $j$ の中継制御に必要なパラメータ群を設定するためのパラメータ設定処理を図3のフローチャートに従って次のように実行する。

【0042】まずノード管理部14は、データ移動部113に対し、トラフィック特性値（推定値）である入力パラメータセット $Ij = \{Rinj, Routj\}$ を与える（ステップ301）。このデータ移動部113には、システムパラメータセット $= \{f1, f2, r1, r2\}$ がノード管理部14により予め与えられているものとする。

【0043】データ移動部113は、データストリーム $j$ の入力の変動分を吸収するためのマージン値 $\alpha1j$ 、及びデータストリーム $j$ の退避・復帰時間の変動分を吸収するためのマージン値 $\alpha2j$ の初期値（固定値）を設定す

14

る（ステップ302）。

【0044】次にデータ移動部113は、データストリーム $j$ 中のデータを退避する際の処理単位となるブロック長 $bj$ の初期値（固定値）を設定する（ステップ303）。

【0045】次にデータ移動部113は、データストリーム $j$ 中のデータを（2次バッファ112への退避なしに）1次バッファ111に蓄積するための領域のサイズ（1次バッファ割り当てサイズ） $Bj$ を入力パラメータセット $Ij = \{Rinj, Routj\}$ 等をもとに決定して設定する処理（ $Bj$ 設定処理）を、図4のフローチャートに従って次のように行う（ステップ304）。

【0046】まずデータ移動部113は、 $Rinj$ 、 $Routj$ の大きさを判定する（ステップ401）。もし、 $Rinj \leq Routj$ ならば、1次バッファ111に溜まるのは送出（再送出）待ち合わせ時間 $wout$ の間の到着データと入力データの変動分によるものであることから、 $Bj$ の値を、例えば $Rinj$ に比例した値（ $k Rinj$ ）に設定する（ステップ402）。ここで $k$ （比例係数）には、 $Bj$ が入力変動を吸収可能となる値を用いればよい。

【0047】一方、 $Rinj > Routj$ ならば、 $Bj$ の値を、例えばサイズ $b$ だけのデータの1次バッファ111へのデータ入力時間 $tini$ と、サイズ $b$ だけのデータが1次バッファ111と2次バッファ112との間を往復（退避・復帰）するのに要する時間（ $w1j + t1j + w2j + t2j + woutj (=0)$ ）と、サイズ $b$ だけのデータの1次バッファ111からのデータ出力時間 $toutj$ との和 $Tj = (tini + w1j + t1j + w2j + t2j + woutj (=0) + toutj)$ の間に再送出するストリーム $j$ のデータ量 $Routj * Tj$ 以上の値（ここでは、時間 $Tj$ の間に再送出するデータ量 $Routj * Tj$ に予め定められた一定量を加えた値）に設定する（ステップ403）。

【0048】このように時間 $Tj$ は、サイズ（ブロック長） $b$ だけのデータの先頭部が1次バッファ111に到達してから、当該データが2次バッファ112との間を（待ち時間なしに）往復し、当該データが1次バッファ111から全て送出されるまでに要する推定時間を表す。したがって、データストリーム $j$ に割り当てる1次バッファ割り当てサイズ $Bj$ の値を、上記のように、時間 $Tj$ の間に再送出するストリーム $j$ のデータ量（ $= Routj * Tj$ ）以上に設定するならば、 $Bj$ の領域がオーバーフローする危険がある場合に、それを回避するためにストリーム $j$ の後続のデータを2次バッファ112に一時退避しても、その退避データを1次バッファ111に復帰させて $Bj$ の領域内の蓄積データに続いて連続的に再送出することが可能となる。即ち再送出の連続性が確保される。

【0049】なお、データ入力時間 $tinj$ 、データ送出時間 $toutj$ は、前記 $tin$ 、 $tout$ の場合と同様にして、 $bj / Rinj$ 、 $bj / Routj$ により求められるもので、

(9)

特開平10-200574

15

それぞれサイズ  $b_j$  だけのデータが到着するのに要する時間の推定値、サイズ  $b_j$  だけのデータを再送出するのに要する時間の推定値である。

【0050】また、推定退避処理時間  $t_{1j}$ 、推定復帰処理時間  $t_{2j}$  は、前記  $t_1$ 、 $t_2$  の場合と同様にして、 $f_1 + b_j / r_1$ 、 $f_2 + b_j / r_2$  により求められる。また、退避待ち時間  $w_{1j}$ 、復帰待ち時間  $w_{2j}$  は、例えば同一の2次バッファ112を利用する全てのデータストリーム  $j$  からのアクセスを入力とした単一処理モデルとして推定可能である。一例として、退避及び復帰の要求は同一優先度で扱われるものとし、それぞれのストリーム  $j$  からの要求発生がランダムで、1ブロックを処理する時間が指数分布に従うとした多元  $M/M/1$  型の近似を行えば、平均待ち時間は以下のように近似表現することができる。

【0051】

$$w_{1j} = w_{2j} = \sum [(t_{1j})^2 * (R_{inj}/b_j) / \{1 - t_{1j} * (R_{inj}/b_j)\} + (t_{2j})^2 * (R_{outj}/b_j) / \{1 - t_{2j} * (R_{outj}/b_j)\}]$$

ここで、 $t_{1j}$ 、 $t_{2j}$  は前記したように  $b_j$  に依存する。したがって上記  $w_{1j}$ 、 $w_{2j}$  は  $b_j$  及び  $R_{inj}$ 、 $R_{outj}$  に依存する。また、 $w_{1j}$ 、 $w_{2j}$  は推定値であることから、以下の説明では、それぞれ推定退避待ち合わせ時間、推定復帰待ち合わせ時間と呼ぶことにする。

【0052】以上のことから、 $R_{inj} > R_{outj}$  の場合の  $B_j$  の値は、 $b_j$  及び  $R_{inj}$ 、 $R_{outj}$  に依存する。データ移動部113は、以上の  $B_j$  設定処理を終了すると、ブロック長  $b$ 、マージン値  $\alpha_{1j}$ 、 $\alpha_{2j}$  等の再設定処理（パラメータ再設定処理）を行う（ステップ305）。このパラメータ再設定処理の詳細については後述する。

【0053】データ移動部113は、ステップ305のパラメータ再設定処理に成功した場合には（ステップ306）、送信ノードから要求されたデータストリーム  $j$  の中継が可能である旨をノード管理部14に通知すると共に、当該データストリーム  $j$  の中継モードを初期モードである退避なしモードに設定する（ステップ307、308）。これに対し、パラメータ再設定処理に失敗した場合には、データ移動部113は送信ノードから要求されたデータストリーム  $j$  の中継が不可能である旨をノード管理部14に通知する（ステップ309）。なお、各データストリームが優先度を持つ場合には、中継中のストリームを含めて、最も優先度の低いストリームの中継を放棄するようにしても構わない。

【0054】ノード管理部14は、データ移動部113からの通知を受けて送信ノードに対して要求されたデータストリーム  $j$  の中継の可否を通知する。次に、データストリーム  $j$  を中継する場合の動作について、データ移動部113による退避・復帰処理を中心に図5及び図6のフローチャートを適宜参照して説明する。

【0055】データ受信部12は、ネットワーク上の送

16

信ノードから送られるデータストリーム  $j$  中のデータ（データパケット）の到着を待つ。一方、データ送信部13は、ノード管理部14からの指示により出力レート  $R_{outj}$  を設定した上で、データストリーム  $j$  中のデータが1次バッファ111に書き込まれるのを待つ。

【0056】データ受信部12は、データパケットが到着したならば、そのパケット中のデータをデータ蓄積部11内の1次バッファ111に書き込む書き込み（データ入力）動作を開始する。一方、データ送信部13は、1次バッファ111へのデータ書き込みを検出して、その先頭から再送処理（データ送出処理）を開始する。

【0057】データ蓄積部11内のデータ移動部113は、データ受信部12により1次バッファ111に書き込まれたデータストリーム  $j$  のデータが当該ストリーム  $j$  について設定したブロック長  $b_j$  を超えると（ステップ501）、退避なしモードである場合には（ステップ502）、ストリーム  $j$  のデータの1次バッファ111内の蓄積量が、データ入力時間（ $b_j$  だけのデータが到着するのに要する推定時間） $t_{inj}$  と1次バッファ111と2次バッファ112との間のデータの往復（退避・復帰）に要する時間（ $w_{1j} + t_{1j} + w_{2j} + t_{2j} + w_{outj}$ （=0））と、1次バッファ111からのデータ送出時間（ $b_j$  だけのデータを再送出するのに要する推定時間） $t_{outj}$  との和  $T_j$  の間に再送出するデータ量を超えたか否かをチェックする（ステップ503）。

【0058】もし、ストリーム  $j$  のデータの1次バッファ111内の蓄積量が上記時間  $T_j$  の間に再送出するデータ量を超えていないならば、データ移動部113は1次バッファ111がオーバーフローする危険性はないものと判断して退避なしモードを継続し（ステップ504）、超えているならば、1次バッファ111のオーバーフローの発生を回避するために、ストリーム  $j$  についての中継モードを退避なしモードから退避ありモードに切り替え設定する（ステップ505）。

【0059】一方、1次バッファ111に書き込まれたデータストリーム  $j$  のデータがブロック長  $b_j$  を超えた場合に（ステップ501）、退避ありモードが設定されているならば（ステップ502）、その先頭から  $b_j$  だけのデータ（データブロック）について1次バッファ111から2次バッファ112への退避処理を開始する（ステップ506）。このとき、退避なしモードから退避ありモードへの切り替え直後の退避処理が開始された場合であれば、1次バッファ111に蓄積されたストリーム  $j$  のデータのうち、退避されない分のデータの再送出がデータ送信部13により続けられる。この退避処理の開始時点においては、 $b_j$  を超えた分のデータが1次バッファ111に書き込まれたデータとされる。

【0060】次にデータ移動部113は、2次バッファ112に退避したデータブロックについて、その先頭部が再送出される時刻を対応するデータストリーム  $j$  の出

(10)

特開平10-200574

17

力レート $R_{outj}$ をもとに推定して、目標復帰完了時刻として設定する（ステップ507）。

【0061】これに対し、1次バッファ111に書き込まれたデータストリームjのデータがブロック長 $b_j$ を超えていない場合には（ステップ501）、退避なしモードであるならば（ステップ508）、ステップ501に戻る。

【0062】また、1次バッファ111に書き込まれたデータストリームjのデータがブロック長 $b_j$ を超えておらず（ステップ501）、しかも退避ありモードであるならば（ステップ508）、データ移動部113は以下の処理を行う。

【0063】まずデータ移動部113は、ストリームjについての退避データブロックについて設定してある目標復帰完了時刻より復帰時間（ $w_{2j} + t_{2j}$ ）だけ遡った時刻が到来したか否かをチェックし（ステップ601）、到来したならば、該当する退避データブロックを2次バッファ112から1次バッファ111へ復帰する復帰処理を開始する（ステップ602）。

【0064】そして、1次バッファ111への復帰が完了したならば（ステップ603）、データ移動部113は、ストリームjのデータの2次バッファ112内の蓄積ブロック数をチェックする（ステップ604）。

【0065】もし、蓄積ブロック数が0であるならば、即ち退避したデータブロックが即刻復帰された場合（これを、即復帰と呼ぶ）には、データ移動部113は、蓄積ブロック数=0の判定（即復帰の判定）が予め定められた一定回数続いたか否かをチェックし（ステップ605）、一定回数続いたならば、ストリームjのバースト的な入力変動が収まって当該ストリームj用に割り当てたサイズ $B_j$ の領域に余裕ができ、オーバフローの危険性はなくなったものとして、ストリームjについての中継モードを退避ありモードから退避なしモードに切り替え設定する（ステップ606）。そしてデータ移動部113はステップ501に戻る。

【0066】一方、目標復帰完了時刻より復帰時間（ $w_{2j} + t_{2j}$ ）だけ遡った時刻が到来していない場合（ステップ601）、蓄積ブロック数が0でない場合（ステップ604）、或いは蓄積ブロック数=0の判定が一定回数続いていない場合（ステップ605）には、そのままステップ501に戻る。

【0067】次に、パラメータ再設定処理（ステップ305）の詳細について、図7及び図8のフローチャートを適宜参照して説明する。データ移動部113は、新規ストリームの追加時、既存ストリームの中継処理の終了時、著しい予測不一致の検出時などにおいて、ブロック長 $b$ 、マージン値 $\alpha_{1j}$ 、 $\alpha_{2j}$ 等のパラメータの再設定処理（更新処理）を次のように行う。なお、著しい予測不一致の検出時とは、送信ノードから通知されたデータストリームjの入力レート $R_{in}$ に対する入力変動が予め定

18

められた許容範囲を超えていることが実測値から検出された場合である（ストリームjの入力レート $R_{in}$ の計測は、当該ストリームjの受信時、データ受信部12にて行われるものとする）。この予測不一致の検出には、データストリームjの転送の突然の終了も含まれる。

【0068】パラメータ再設定処理では、まずデータ移動部113は、各ストリームjの最新のトラフィック特性値 $R_{inj}$ 、 $R_{outj}$ 等に基づき、1次バッファ利用率 $\rho_1$ 及び2次バッファアクセスチャネル利用率 $\rho_2$ を求める（ステップ701、702）。

【0069】ここで1次バッファ利用率 $\rho_1$ の値は、各退避なしストリームj（退避なしモードにあるストリームj）の各々に対して割り当てられている1次バッファ111内の領域のサイズ（1次バッファ割り当てサイズ） $B_j$ の（マージン値 $\alpha_{1j}$ を考慮した）総和 $\Sigma 1$ （即ち全ての退避なしストリームが1次バッファ111上で占めると予想される全体サイズ）と、各退避ありストリームjに関して、1次バッファ111内の入力バッファ領域長 $L_{inj}$ と出力バッファ領域長 $L_{outj}$ の和（ $L_{inj} + L_{outj}$ ）に、1次バッファ111内の退避バッファ長 $b_j$ と復帰バッファ長（=退避ブロック長） $b_j$ に各々の推定利用時間の割合を掛け合わせた値を加えたものの（マージン値 $\alpha_{2j}$ を考慮した）総和 $\Sigma 2$ （即ち全ての退避ありストリームが1次バッファ111上で占めると予想される全体サイズ）とを足し合わせた値を、1次バッファ111のサイズ（1次バッファ量） $L_1$ で割った値であり、

$$\rho_1 = (\Sigma 1 + \Sigma 2) / L_1$$

のように表される。

【0070】ここで、 $\Sigma 1 = \Sigma \{B_j * (1 + \alpha_{1j})\}$ である。

【0071】また、推定利用時間の割合は、退避バッファ長（=退避ブロック長） $b_j$ については、ストリームjのデータがサイズ $b$ だけ到着するのに要する時間（サイズ $b$ のデータの入力時間）に対する、サイズ $b$ のデータ（データブロック）を1次バッファ111から2次バッファ112に退避するのに要する時間（復帰時間）の割合、即ち $(w_{1j} + t_{1j}) / t_{inj}$ であり、復帰バッファ長（=退避ブロック長） $b_j$ については、ストリームjのデータをサイズ $b$ だけ再送出するのに要する時間（サイズ $b$ のデータの出力時間）に対する、サイズ $b$ のデータ（データブロック）が2次バッファ112から1次バッファ111に復帰するのに要する時間（復帰時間）の割合、即ち $(w_{2j} + t_{2j}) / t_{outj}$ である。

【0072】したがって $\Sigma 2$ は、 $\Sigma 2 = \Sigma \{ (L_{inj} + L_{outj} + b_j * (w_{1j} + t_{1j}) / t_{inj} + b_j * (w_{2j} + t_{2j}) / t_{outj} ) * (1 + \alpha_{2j}) \}$ のように表される。

【0073】また2次バッファアクセスチャネル利用率

(11)

特開平10-200574

19

20

$\rho 2$  の値は、退避ありストリーム  $j$  に関する退避データ  
トラフィック  $t1j/tin j$  と復帰データトラフィック  $t$   
 $2j/tout j$  との和の総和であり、

$$\rho 2 = \Sigma (t1j/tin j + t2j/tout j)$$

のように表される。

【0074】上記  $\rho 1$ 、 $\rho 2$  の値は、各ストリーム  $j$  に  
与えられた1次バッファ割り当てサイズ  $Bj$  及びブロッ  
ク長  $bj$  の値により変化する。データ移動部113は、  
 $\rho 1$ 、 $\rho 2$  を算出すると、 $\rho 1$  及び  $\rho 2$  と（容量限界を  
意味する）1とを比較して、 $\rho 1$  及び  $\rho 2$  が共に1以下  
であるか否かを判定し（ステップ703）、共に1以下  
であるならば、現在各ストリーム  $j$  に与えている  $Bj$  及  
び  $bj$  の値は適性であると判断し、パラメータ再設定処  
理を終了する。

【0075】これに対し、 $\rho 2$  は1以下であるものの、  
 $\rho 1$  が1を超えているならば（ステップ704）、デー  
タ移動部113はまず、1つのストリーム  $j$  を選択し、  
そのストリーム  $j$  に割り当てられているブロック長（退  
避ブロック長） $bj$  を予め定められた一定値  $\Delta b$  だけ小  
さくした場合の  $\rho 1$  及び  $\rho 2$  の値を（上記ステップ70  
1、702と同様にして）再計算する（ステップ70  
5、706）。ここで、 $bj = bj - \Delta b$  の操作の対象  
には任意のストリームを選択しても構わないが、本処理  
が新規ストリームの追加、或いは著しい予測不一致の検  
出によって起動された場合には、その要因となったスト  
リームを選択するとよい。

【0076】もし、ステップ706で再計算した  $\rho 1$  及  
び  $\rho 2$  の値が共に1以下であるならば（ステップ70  
7）、データ移動部113はその際の  $bj$  を該当するス  
トリーム  $j$  用に採用し、パラメータ再設定処理を終了  
する。一方、再計算した  $\rho 1$  及び  $\rho 2$  のうちの少なくとも  
一方が1を超えているならば、全ての退避ありストリー  
ム  $j$  について、そのブロック長（退避ブロック長） $bj$   
を  $\Delta b$  だけ小さくした場合の  $\rho 1$  及び  $\rho 2$  の値を再計算  
する（ステップ708、709）。

【0077】ここで、ステップ709で再計算した  $\rho 1$   
及び  $\rho 2$  の値が共に1以下であるならば（ステップ71  
0）、データ移動部113はその際の各  $bj$  を該当する  
ストリーム  $j$  用に採用し、パラメータ再設定処理を終了  
する。これに対し、ステップ709での再計算を行って  
も  $\rho 1$  及び  $\rho 2$  のうちの少なくとも一方が1を超えてい  
るならば、後述するマージン軽減処理（ステップ71  
1）に移る。

【0078】なお、上記ステップ708において1つの  
ストリーム  $j$  について、そのブロック長（退避ブロック  
長） $bj$  を  $\Delta b$  だけ小さくして、ステップ709で  $\rho 1$   
及び  $\rho 2$  を再計算し、その再計算後の  $\rho 1$ 、 $\rho 2$  の値の  
少なくとも一方が依然として1を超えている場合には、  
 $bj$  を  $\Delta b$  だけ小さくするストリーム  $j$  を切り替えて、  
 $\rho 1$ 、 $\rho 2$  の再計算と評価を、一定回数を超えない範囲

で、 $\rho 1$ 、 $\rho 2$  の値が共に1以下となるか、或いは対象  
となるストリーム  $j$  がなくなるまで繰り返すようにして  
も構わない。

【0079】一方、ステップ701、702で算出され  
た  $\rho 1$ 、 $\rho 2$  のうち、 $\rho 1$  は1以下であるものの、 $\rho 2$   
が1を超えているならば（ステップ703、704、8  
01）、データ移動部113はまず、1つのストリーム  
 $j$  を選択し、そのストリーム  $j$  に割り当てられているブ  
ロック長（退避ブロック長） $bj$  を  $\Delta b$  だけ大きくした  
場合の  $\rho 1$  及び  $\rho 2$  の値を（上記ステップ701、70  
2と同様にして）再計算する（ステップ802、80  
3）。

【0080】もし、ステップ803で再計算した  $\rho 1$  及  
び  $\rho 2$  の値が共に1以下であるならば（ステップ80  
4）、データ移動部113はその際の  $bj$  を該当するス  
トリーム  $j$  用に採用し、パラメータ再設定処理を終了  
する。一方、再計算した  $\rho 1$  及び  $\rho 2$  のうちの少なくと  
も一方が1を超えているならば、全ての退避ありストリー  
ム  $j$  について、そのブロック長（退避ブロック長） $bj$   
を  $\Delta b$  だけ大きくした場合の  $\rho 1$  及び  $\rho 2$  の値を再計算  
する（ステップ805、806）。

【0081】ここで、ステップ806で再計算した  $\rho 1$   
及び  $\rho 2$  の値が共に1以下であるならば（ステップ80  
7）、データ移動部113はその際の各  $bj$  を該当する  
ストリーム  $j$  用に採用し、パラメータ再設定処理を終了  
する。これに対し、ステップ807での再計算を行って  
も  $\rho 1$  及び  $\rho 2$  のうちの少なくとも一方が1を超えてい  
るならば、後述するストリーム中継棄却処理（ステップ  
808）に移る。

【0082】また、ステップ701、702で算出され  
た  $\rho 1$ 、 $\rho 2$  が共に1を超えている場合には（ステップ  
703、704、801）、データ移動部113はステ  
ップ711のマージン軽減処理に移る。

【0083】次に、上記ステップ711のマージン軽減  
処理について図9のフローチャートを参照して説明す  
る。まずデータ移動部113は、各ストリーム  $j$  の退避  
・復帰時間の変動分を吸収するために割り当てたマージ  
ン値  $\alpha 2j$ （通常は1未満の正の小数值）が正である範囲  
で、当該  $\alpha 2j$  の値を予め定められた一定値  $\Delta \alpha$  だけ減じ  
た後（ステップ901、902）、 $\rho 1$  の値を（上記ス  
テップ701と同様にして）再計算する（ステップ90  
3）。

【0084】もし、ステップ902で再計算した  $\rho 1$  の  
値が1以下となったならば（ステップ904）、デー  
タ移動部113はその際の各  $\alpha 2j$  を該当するストリーム  $j$   
用に採用し、パラメータ再設定処理を終了する。

【0085】これに対し、ステップ902で再計算した  
 $\rho 1$  の値が依然として1を超えているならば、データ移  
動部113は上記ステップ901以降の処理を再度実行  
する。そして各  $\alpha 2j$  を0にして（ステップ905）、 $\rho$

50

(12)

特開平10-200574

21

1を再計算しても(ステップ906)、 $\rho 1$ の値が1以下とならないならば(ステップ907)、データ移動部113は、各ストリームjの入力の変動分を吸収するために割り当てたマージン値 $\alpha 1j$ (通常は1未満の正の小数值)が正である範囲で、当該 $\alpha 1j$ の値を $\Delta \alpha$ だけ減じた後(ステップ908, 909)、 $\rho 1$ の値を(上記ステップ701と同様に)再計算する(ステップ910)。

【0086】もし、ステップ910で再計算した $\rho 1$ の値が1以下となったならば(ステップ911)、データ移動部113はその際の各 $\alpha 1j$ ,  $\alpha 2j$ を該当するストリームj用に採用し、パラメータ再設定処理を終了する。

【0087】これに対し、ステップ910で再計算した $\rho 1$ の値が依然として1を超えているならば、データ移動部113は上記ステップ908以降の処理を再度実行する。そして各 $\alpha 1j$ を0にして(ステップ912)、 $\rho 1$ を再計算しても(ステップ913)、 $\rho 1$ の値が1以下とならないならば(ステップ914)、データ移動部113は、図8中のステップ808のストリーム中継棄却処理に移る。

【0088】次に、上記ステップ808のストリーム中継棄却処理について図10のフローチャートを参照して説明する。まずデータ移動部113は、中継棄却の対象となるストリームを選択し、その選択したストリームの中継処理を放棄する(ステップA1)。このステップA1では、新規ストリームの追加時(新たにストリームの中継が要求された場合)であれば、その新規ストリームか、中継中のストリームを含めて、最も優先度の低いストリームを選択すればよい。また、新規ストリームの追加時以外であれば、中継中のストリームの中で最も優先度の低いストリームを選択すればよい。

【0089】次にデータ移動部113は、図7及び図8のフローチャートに示すパラメータ再設定処理を呼び出して、残りのストリームjについて、 $\rho 1$ ,  $\rho 2$ の計算等を行うことでブロック長 $b_j$ 等のパラメータを再設定し(ステップA2)、中継処理を継続する(再度のフレーム中継棄却処理が不要の場合)。

【0090】以上に述べた実施形態では、ストリームjについての1次バッファ111内のデータ蓄積量が時間Tjの間に再送出するデータ量を超えた場合には、直ちに退避モードに切り替えるものとして説明したが、例えばその差分以上の空き領域(非割り当て領域)が1次バッファ111上に存在する場合には、その差分、或いはその差分に予め定められた一定量を加えた値だけBjを増やして退避なしモードを継続し、それでも不足する場合には、退避モードに切り替えるようにしても構わない。

【0091】なお、図1の構成の中継ノードの各部の処理機能、即ちデータ受信部12、データ送信部13、ノード管理部14、及びデータ蓄積部11内のデータ移動部113の処理機能は、その処理機能をCPUを備えた

22

コンピュータ等の情報処理機器により実現させるためのプログラムを記録した(メモ리카ード、フロッピーディスク、CD-ROM等の)記録媒体を用い、当該記録媒体を情報処理機器に装着して、当該記録媒体に記録されているプログラムを情報処理機器で読み取り実行させることによっても実現される。

【0092】

【発明の効果】以上詳述したように本発明によれば、中継ノードにおける一時蓄積バッファを高速メモリによる1次バッファ手段と低速大容量の2次バッファ手段との組み合わせによって構成し、1次及び2次バッファ手段間の適切なデータ移動及び1次バッファからの適切なデータ送出を制御することにより、比較的低価格でありながら、ストリーム中継に伴う遅延を小さくできると共に、大容量のバースト的なデータをオーバーフローの発生を招くことなく一時蓄積することができる。

【図面の簡単な説明】

【図1】本発明の一実施形態に係る中継ノードの構成を示すブロック図。

【図2】図1中のデータ蓄積部11内でのバッファ111, 112間のデータ移動制御の手順の概略を説明するためのタイミングチャート。

【図3】同実施形態におけるパラメータ設定処理を説明するためのフローチャート。

【図4】同実施形態におけるBj設定処理を説明するためのフローチャート。

【図5】同実施形態における中継動作について、データ移動部113による退避・復帰処理を中心に説明するためのフローチャートの一部を示す図。

【図6】同実施形態における中継動作について、データ移動部113による退避・復帰処理を中心に説明するためのフローチャートの残りを示す図。

【図7】同実施形態におけるパラメータ再設定処理を説明するためのフローチャートの一部を示す図。

【図8】同実施形態におけるパラメータ再設定処理を説明するためのフローチャートの残りを示す図。

【図9】図7中のステップ711のマージン軽減処理を説明するためのフローチャート。

【図10】図8中のステップ808のフレーム中継棄却処理を説明するためのフローチャート。

【図11】従来の中継ノードの構成を示すブロック図。

【符号の説明】

10…中継ノード  
11…データ蓄積部  
12…データ受信部  
13…データ送信部  
14…ノード管理部  
111…1次バッファ  
112…2次バッファ  
113…データ移動部(パラメータ設定手段、モード設

(13)

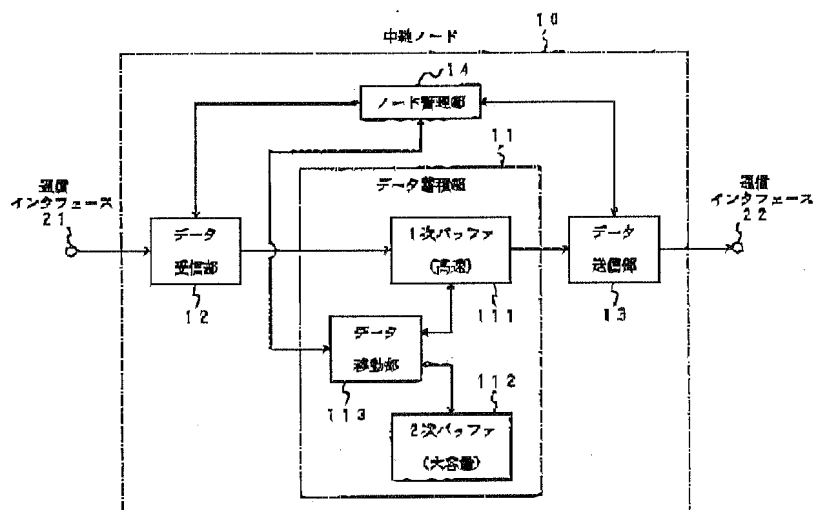
特開平10-200574

23

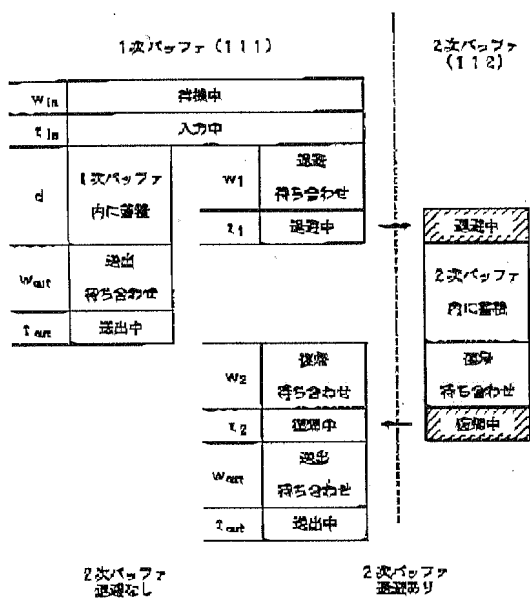
24

定手段、退避手段、復帰手段)

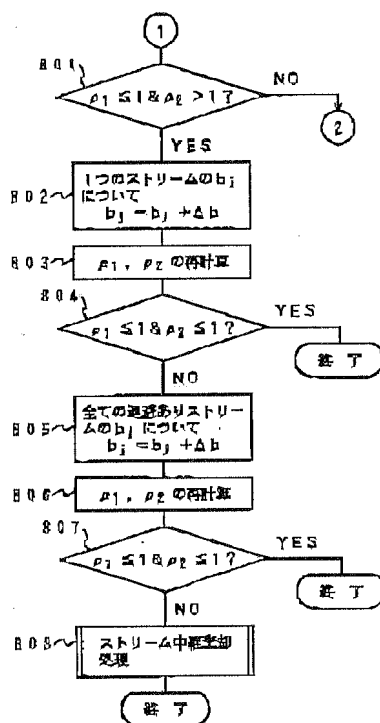
【図1】



【図2】



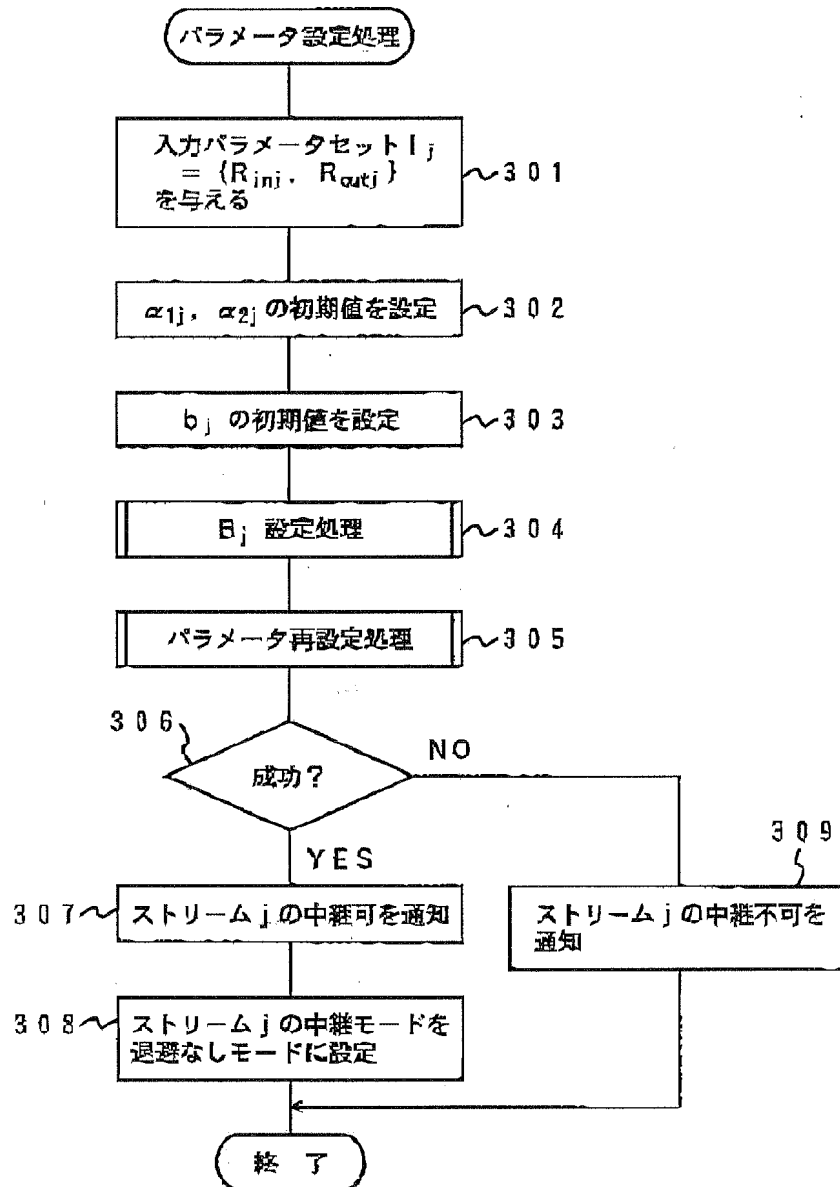
【図8】



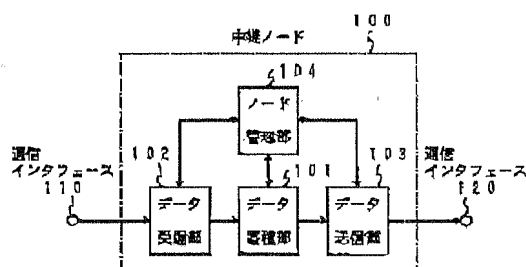
(14)

特開平10-200574

【図3】



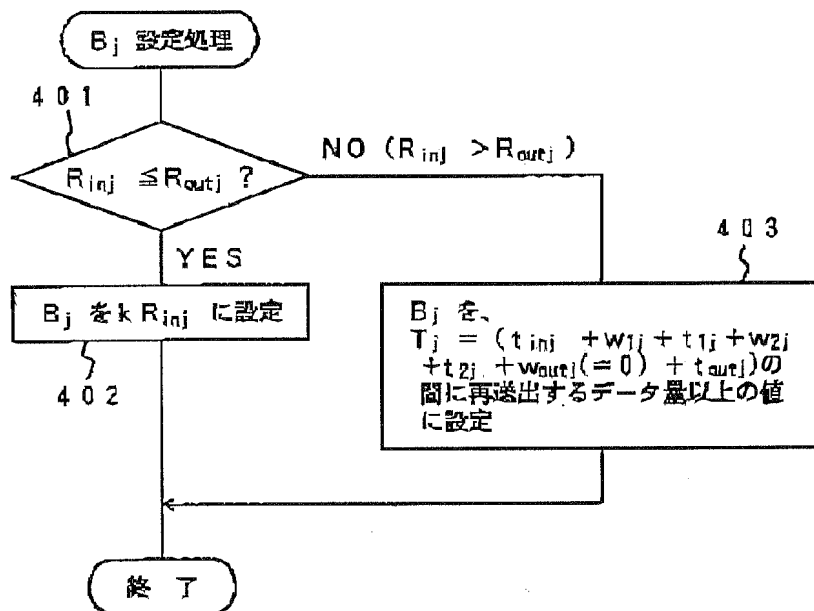
【図11】



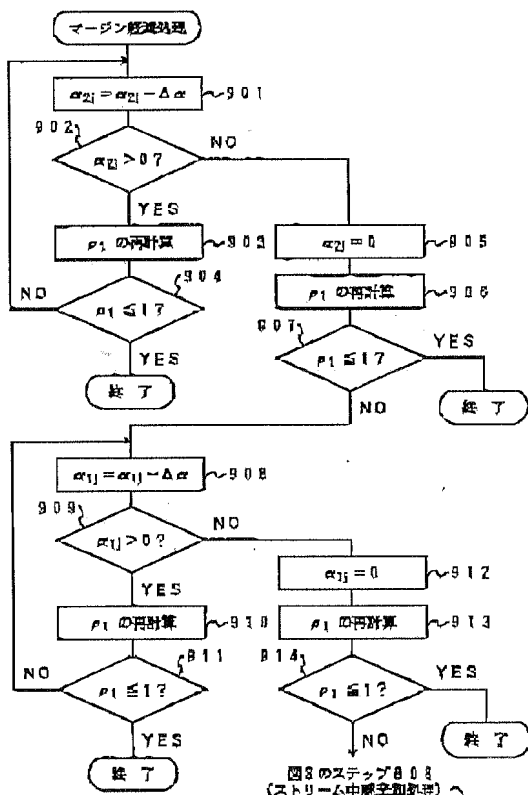
(15)

特開平10-200574

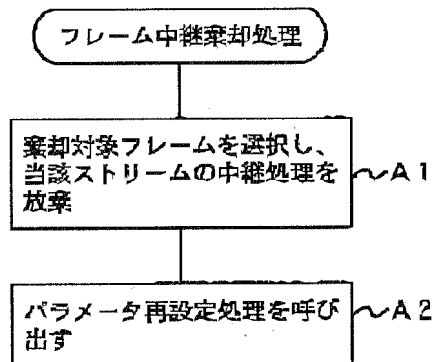
【図4】



【図9】



【図10】

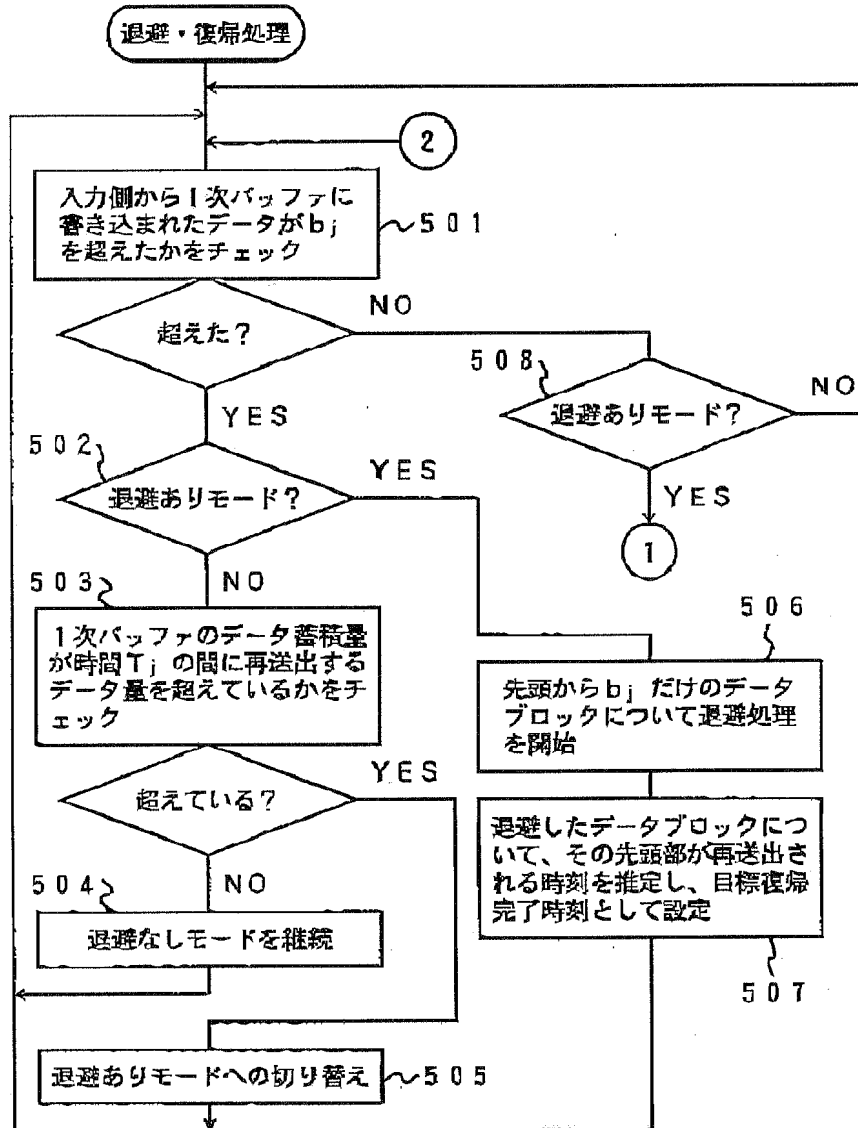




(16)

特開平10-200574

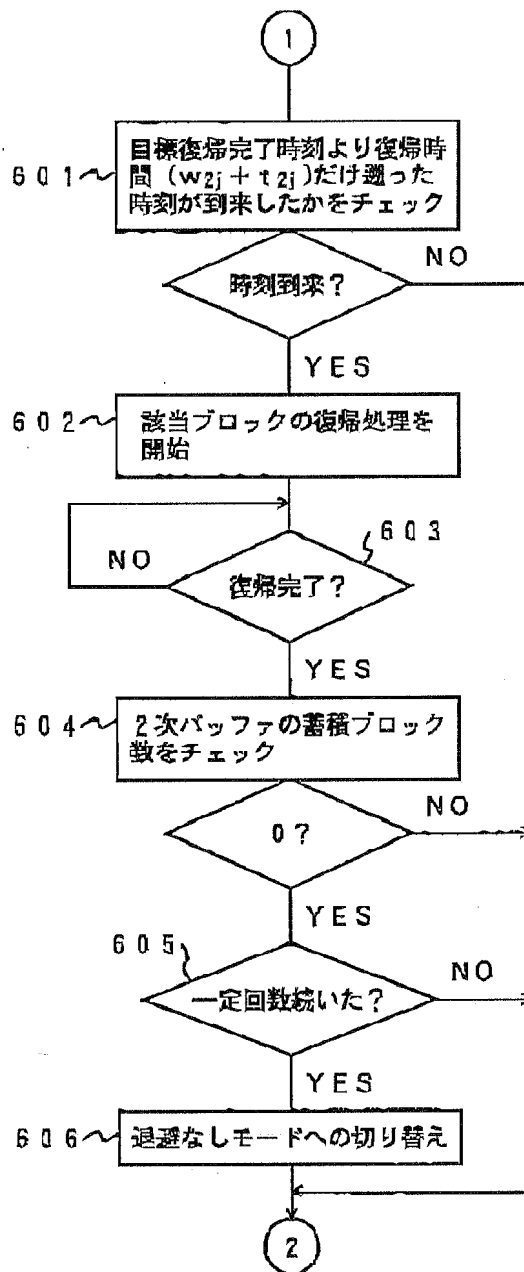
【図5】



(17)

特開平10-200574

【図6】



(18)

特開平10-200574

【図7】

